Captioning Effects on Television News Learning.

Subjects, 100 undergraduate journalism students, viewed news stories in a quiet classroom. They first viewed a dummy story followed by a brief distractor test meant to purge short-term memory and prevent rehearsal of story information. Subjects then viewed actual news stories of two types, concrete/picture stories and abstract/word stories, in which the caption variable was manipulated. For each news story, subjects identified whether freeze frames they were shown had appeared in the dummy story and answered multiple choice questions (testing both recall and understanding) about the content. The process was repeated until five more stories had been viewed. Results indicated that captions did improve recall for the abstract/word stories, but impeded understanding in the concrete/picture stories. Captions may have impeded understanding of the picture stories by distracting attention from the visuals. On the other hand, because word stories do not lend themselves to reinforcing visualization, cues like captions serve as reinforcements. This implies that news producers would do best to omit captions from stories with strong visual content and use caption graphics to clarify abstract, word oriented stories. (References, notes, and two tables conclude the study.) (SKC)
CAPTIONING EFFECTS ON TELEVISION NEWS LEARNING

Stephen D. Reese

William R. Davie

Department of Journalism, College of Communication
University of Texas, Austin 78712
(512) 471-7708/1966


"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stephen Reese

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."
Captioning Effects on Television News Learning

Over the last decade, many studies have posed the question: how much of television news does the audience absorb? The findings show that viewers often remember either a distorted version of the message, or little at all (e.g., Katz, Adoni & Parness, 1977; Neumann, 1976; Robinson, Davis, Sahin & O'Toole, 1980; Sahin, Davis & Robinson, 1981; and Stauffer, Frost & Rybalt, 1978). To address this concern, researchers have turned to examining message structure and format characteristics to identify those which best facilitate learning.

Visual format features, in particular, have been found to affect the amount and kind of information that news viewers learn (reviewed in Berry, Gunter & Clifford, 1981; Gunter, 1980; and Gunter, Berry & Clifford, 1982). Visual illustrations, for example, have been found most effective when they reinforce, or are "redundant" with, the script (Drew & Grimes, in press; Reese, 1984; and Son, Reese & Davie, 1986).

In addition to non-verbal illustration, however, the use of "visual-verbal" print material has become common. Captions no longer are used to simply identify newsmakers with "supers," but are also widely used in titling graphics and enumerating important points in a script. However, this form of verbal reinforcement has received little research attention, and its impact on learning remains unclear. The purpose of this study is to examine the extent to which captioning assists viewers in learning more from television news stories.

Although many multiple-channel communication studies have tested memory for lists of words and pictures (reviewed in Hartman, 1961), a news story presents a different task than simply recalling lists of information. Stories have internal structure, and viewing them requires attending to certain parts more than others, remembering details, and trying to make sense of them. Studies examining news stories and captions have provided mixed results.

Continuous captions have been found to either impede or have no effect on learning from news stories (Reese, 1984). This may be explained by noting the limitations of single-channel processing, a model which describes humans as capable of attending to only one channel at a time. Switching attention from one channel (spoken text) to another (written text) impairs performance when processing demands are heavy because some information is lost in the process (Broadbent, 1958; Treisman, 1969; and Treisman & Davies, 1973).

However, if processing demands are reduced via non-continuous captions, they should aid viewers by reinforcing important material. Findahl and Hoijer (1976, 1981), for example,
show that, when used intermittently in a story, captions do benefit viewers by directing attention, and cuing recall of specific content. In the Findahl and Hoijer studies, captions were considered simply another means (along with visualization and repetition) of emphasizing certain story elements over others, bringing them to center stage, as it were. Emphasizing contextual detail (event causes and consequences) was said to have helped viewers remember the details of the event itself as well.

Captions should also be beneficial as an information processing aid. Encoding spoken (acoustic) information represents a low effort psychological task, which makes it more difficult to develop associative semantic links. Reading, on the other hand, requires more effortful semantic coding. As a result, Kellerman (1985) suggests that "... increased use of verbal representations presented visually could promote use of semantic codes (p. 112)" (and, thus, facilitate memory).

This study extends the Reese (1984) study by examining the effect of captions using stories with film accompaniment, the industry norm, rather than still illustrations used by Findahl and Hoijer. In addition, a more realistic captioning manipulation is employed. Rather than superimposing a continuous transcript of the story, key word captions are used intermittently to emphasize story material.

In addition, this study examines the relative benefit of captions in different types of stories. A distinction is made between stories based on the nature of the event they portray. One class of stories deals with concrete, easy to visualize events, such as fires, earthquakes, and public protests. The other type describes abstractions, is more difficult to visualize, and relies more on words to get the point(s) across. Stories about legislation, court decisions, and negotiations fit into this category. The visuals in these so-called "word" stories do not, by definition, have as much to do--are not as redundant with--the script. Photo opportunity shots of President Reagan greeting a Russian ambassador, for example, relate only marginally to the negotiations they held. If captions help direct attention and encourage semantic processing, they should be particularly beneficial in these stories, which don't benefit as much from visual reinforcement. Their visuals, in fact, may often distract viewers from the main point of the story.
CAPTIONING EFFECTS

METHODS

March 31, 1987

Data were gathered in the late Spring and early Summer semesters of 1986, using undergraduate subjects at the University of ___ enrolled in various courses in the Department of Journalism. The experiment employed a 2 x 2 factorial design. The captioning manipulation represented one factor. The other factor was type of story, either picture or word.

Procedure

Subjects (N=100) were asked to sign up for one of several possible viewing sessions (four to six subjects in each). These were then randomly assigned to the four experimental conditions. No intact classes were used. This procedure (similar to that used in Reese, 1984; and Son, Reese and Davie, 1986) approximates random assignment by keeping viewing sessions small. Each viewing session was held in a quiet classroom. The subjects were instructed that they were to participate in a study of what people learn from television news programs and that they would be shown five different television news stories. They were told to just watch and listen to the stories as they normally would at home.

They were first shown a dummy story (Louisiana Governor), which was used to serve as a check on randomization. Following that story they were instructed to turn to the first section of the questionnaire. A one-minute distractor test was administered to purge short-term memory and prevent rehearsal of story information (as in Gunter, Berry and Clifford, 1981).2

The visual recall test was then administered (similar to Drew & Grimes, 1986). Subjects were shown a series of seven successive visuals, or "freeze frames," on videotape, some taken from the preceding story, some from stories subjects had not seen (about 4 seconds each in duration). As they saw each visual, they were to mark on their questionnaire whether or not the visual had been in the story they just saw. Immediately following this test, subjects were asked about the central point of the story, followed by the four immediate recall multiple choice questions. After seeing another story the process was repeated until they had seen all five stories. Following a two-minute session of six more puzzles, subjects were again shown the visuals from all five stories to identify. They then completed the delayed recall multiple choice items from all five stories (the questions from the Louisiana Governor dummy story were simply repeated). The entire session lasted about 45 minutes.
Development of Stimulus Materials

Story selection. Several network newscasts were videotaped in September of 1985. Twelve stories were selected to be pretested: six were thought a priori to be good examples of "concrete/picture" stories, and six were thought to be "abstract/word" stories. In order to select the best examples and confirm this conceptual distinction, these stories were subjected to a validity check. They were shown in succession to three judges (journalism graduate students), who were also given a complete transcript of each report. After viewing the stories, they were instructed to rank the stories from most concrete to most abstract. The major distinction to be made was between stories about physical events, either recorded or reconstructed (concrete), and non-visualizable processes (abstract).

Following this procedure, eight stories were selected (four of each type) which produced the optimum intercoder reliability, C.R.=.93 (Holsti, 1969). To compute this agreement, each judge's ranking was converted into a series of 25 separate decisions, representing each possible discrimination between the five word and five picture stories. The "correct" judgments would have ranked each picture story as more concrete than any of the respective word stories. The reported coefficient is the average percentage of agreement by the three judges across these 25 decisions.

The four concrete/picture stories selected (and their duration) were reports on Hurricane Gloria (1:34), the Mexico City Earthquake (1:40), the Delta airliner crash in Dallas (1:32), and a terrorist attack in Cyprus (1:10). The four abstract/word stories concerned AIDS (1:30), the Rainbow Warrior scandal in France (1:17), the Geneva Arms talks (1:16), and South Africa unrest (1:40). A story about the trial of Louisiana Governor Edwin Edwards (1:27) was selected to serve as an unmanipulated dummy story.

Captioning manipulation. Some of the selected stories already contained some captioning to illustrate the main points (e.g., in the AIDS story, the drugs holding the most promise were "Suramin, Forscarinet, Compound 5, Riboviran and HPA-23," presented in sequence on the screen as the reporter read the list). Additional intermittent captions were added to each story so as to reinforce all the major details (about seven captions in each story). These were similar to caption graphics used in many news stories, brief summaries of the points in the script, left on-screen just long enough to be read easily.
Measures

Given that recall and understanding are distinct processes (e.g., Woodall, Davis & Sahin, 1983), viewers may recall facts from a story without having assimilated them into a larger framework of understanding, they are measured separately below (as in Son, Reese & Davie, 1986).

Recall. Questions were derived from story transcripts, particularly those blocks of copy which were captioned. Ten questions from each story were pretested. Following an item analysis the two items in each group of 10 showing the least reliability were dropped. In each resulting group of eight, four questions were selected randomly to test immediate recall and four were selected to tap delayed recall. Additional item analysis was conducted on each block of four questions, and in all but a few blocks an item was dropped to further refine reliability.

Understanding. The ability of subjects to reproduce the central point of the stories was used to operationalize understanding. Following a technique used by Sahin, Davis and Robinson (1981), the central points were defined as the essential element of each story, which a reporter would have wanted a viewer to gain from the story (listed below). Subjects' open-ended responses were coded by the authors according to their correspondence with those points. The coding of a subset of responses was compared and a reliable agreement achieved (C.R.=.80).

1. Hurricane Gloria: South Carolina and Virginia residents awaited the arrival of the worst storm ever to threaten the east coast, a storm which could have threatened states as far as Maine.

2. Earthquake: Foreign rescue teams were searching for Mexico City earthquake survivors, but were encountering many obstacles.

3. Terrorists: Gunmen had seized an Israeli yacht in Cyprus, killing three on board before surrendering to Cypriot and Egyptian authorities.

4. Delta 191: Transcripts of the final minutes before the Delta crash in Dallas showed that the pilots saw the thunderstorms but could not fight the effects of wind shear.
5. AIDS: Scientists were testing hundreds of drugs designed to inhibit the AIDS virus, which initially infects the brain before migrating to the rest of the body.

6. Arms Control: The Soviet foreign minister had outlined a new arms control proposal to be presented in Geneva, although the administration insisted its "Star Wars" defense plan would continue.

7. Rainbow Warrior: The French prime minister had blamed the former defense minister and secret service chief in the sinking of the Greenpeace "Rainbow Warrior" ship.

8. South Africa: President Botha had presented a plan for Apartheid reform, responding to increasing pressure from the business community, hurting from a black boycott.

**Visual Recall.** Visual recall was the number of correct discriminations subjects made when shown a set of freeze-frame images (as described above). The highest score was achieved by correctly determining whether each visual was or was not in the story just seen.

**RESULTS**

ANOVA was used to detect any significant differences among the four groups on the dependent variables. Examining total (across all stories) combined (immediate and delayed) recall showed a main effect for the captioning factor (F[1, 96]=4.20, p=.043), but none for story type (F=1.59, p=.21) or the interaction between the two factors (F=1.07, p=.31). A similar analysis of total recall showed a significant main effect for captioning on immediate recall (F=3.96, p=.049) but not on delayed recall (F=2.49, p=.118). So, overall, captioning had a positive impact on recall.

Next, differences were examined within the two story types. As seen in Tables 1 and 2 captions enhanced recall from the word stories while they did not appear to do so in the picture stories. Although this discrepancy was not strong enough to produce a significant interaction effect in the ANOVA results, t-test comparisons show three significant differences (below the p=.05 level) among the immediate and delayed recall measures for the word stories, but no significant differences among the picture stories. The total immediate and delayed recall scores show this more clearly. Among those subjects viewing the word stories, those seeing the captioned versions showed greater immediate, delayed and total recall than those seeing the
Captioning had no significant overall main effect on understanding. However, when the story types were examined separately significant comparisons did emerge. Captioning did not significantly affect understanding of the word stories, but in two of the three picture stories (Hurricane Gloria and Earthquake) it actually impeded understanding (See Table 2). The difference in the Terrorist story was not significant but in a direction consistent with the other two.

As seen in Tables 1 and 2, memory for story visuals was uniformly high. However, ANOVA results did show significant differences. A main effect for story type \((F=4.11, p=.045)\) showed that the picture stories were at least visually more memorable. Although there was no significant captions effect \((F=.73, p=.40)\) the interaction between story type and captioning was significant \((F=3.99, p=.049)\). Although the differences are not great, the means shown in the tables show that captions tend to wash out the natural differences between picture and word stories.\(^8\)

DISCUSSION

Captions did improve recall for the abstract/word stories, but they did not seem to make much difference among the concrete/picture stories. In fact, captions impeded viewer understanding of the picture stories. The beneficial captioning effect on recall suggests that making them intermittent, vs. continuous, lowers the processing demands on viewers and permits the captions to highlight key details rather than serve as a distraction. In the word stories, the visual illustration could not, by definition, illustrate abstract concepts, as they could in the graphic, picture stories. Therefore, captions in these reports provided an additional source of redundancy and cues to help override the potentially distracting visuals.

The different impact of captions on understanding and recall points out the importance of examining them separately. As seen here, captions can benefit recall but not understanding, as in the word stories, and, while having no effect on recall from picture stories, they impeded understanding of two of those three reports. The results for understanding are consistent with those reported in Reese (1984), which showed lower understanding (measured the same way) of a captioned version of a story about El Salvador guerrillas (which this study would have defined as a picture story).
Captions may have impeded understanding of the picture stories by distracting attention from the visuals. Captioning reduced visual recall from the picture stories, suggesting that a distraction process did occur. This confirms what may be considered intuitively obvious: pictures are important in helping viewers understand essentially picture stories. Captions may work against understanding these picture stories by highlighting details throughout a story. This may obscure the real gist, or main point, of the report, leading viewers to assimilate the entire story, as if each detail was as important as the next. This may prevent a more holistic, integrated comprehension. A more judicious highlighting of critical points would seem more appropriate. (Note that the visual recall measure may more properly be called a recognition task, which produced high levels of performance. A more difficult recall measure would perhaps show clearer effects of format features on visual memory).

For example, Findahl and Hoijer captioned consequences of an event in a story in order to help recall. Focusing on these contextual details was more helpful than emphasizing the event per se. Similarly, Son et al. (1987) found that recaps of a story's central point helped understanding by, presumably, helping viewers isolate the essential gist of the story.

A deliberate decision was made in this study to "teach" subjects that they were to learn from the stories (by testing them after the first dummy story). They were tested after each story so as to compare learning across stories without worrying about a primacy/recency effect. The design, along with the subject pool (ostensibly news-interested journalism students) and the use of several high-profile stories may have mitigated against greater observed differences. Message structure features like captions may be particularly beneficial to the less informed or more passive viewers, giving them additional cues in processing unfamiliar stories, or those to which they give only scant attention.

At any rate, it's encouraging that captions aid learning from at least some types of stories, given the low level of viewer recall discussed earlier. Word stories, by definition, do not lend themselves to reinforcing visualization, so additional cues like captions must serve instead. Although news producers intuitively have used such caption graphics for some time to clarify stories, this empirical test helps confirm their effectiveness.
REFERENCES


NOTES

1. The volunteer subjects came primarily from a large introductory communication theory course, but also from introductory newswriting and public relations classes. The subjects were all juniors and seniors, and no broadcast journalism students were used.

2. Each distractor quiz consisted of a series of brain-teaser "pundles." For example, the first pundle, "ONCE / 4:00 p.m." was revealed as "once upon a time." Four such "pundles" were posed after each subsequent news story viewing.

3. The judges instructions were as follows: "You have been given the transcripts of the stories shown on the videotape. Use these to help you make the rankings described below. Note that the soundbite interviews in the story transcriptions have been omitted. As you know, some stories may be considered word stories and some picture stories (some obviously have both characteristics). This is another way of saying they are abstract (word stories) or concrete (picture stories). The concrete/physical story is easy to visualize. The central point of the story is some physical event which the viewer may see on the screen. The pictures are important to help explain the story. Note that it is not always possible for the camera to witness the event, but it may be visually "reconstructed" after the fact. Think of what the central point of each story is. Is it a physical event either observed or observable (had the camera been present)? If so the story is more like a concrete/picture story. By contrast, abstract/word stories are more difficult to visualize. The central point is not what happened as a physical event, but what ideas were discussed, what measures were passed or what controversy was portrayed. Many political stories and judicial stories fall in this category, and the pictures simply serve as a backdrop to the reporter's script. Note that these stories may feature an actual event (meetings, arrivals, etc.), but that is often not the central point of the story. After you have seen the stories, take the story scripts and arrange them in order of concreteness/abstractness. That is, place the story you feel is most concrete at the top of the stack and number it "1." Place the story you feel is the next most concrete second in the stack and label it "2." Continue this process until you have arrayed the stories in the order you feel is most appropriate.
CAPTIONING EFFECTS

Each story will be ranked from 1 to 12. The story last in the stack, number 12, should be the one you feel is most abstract. You may have more difficulty differentiating between the middle stories, but just make the choice you feel is best. Be sure to clearly number the stories at the top of each script, based on your ranking. Please don't worry about whether or not your ranking corresponds to that of other judges.

4. In the Hurricane Gloria story, for example, the reporter's words, "there have been only four hurricanes this century as powerful," became a caption, "Only 4 this century as powerful." The captions were added so as to fit the story, either on the bottom third of the screen or using the full screen as needed. The original visuals were retained as a backdrop. In the non-captioned version, those stories which originally contained some captions were altered to remove them, replacing them with appropriate video. Although some stories were shortened for the experiment, the original structure was retained, including reporter "stand-ups." However, these on-camera remarks were not captioned, nor were they included in the recall questions.

5. Students in a large introductory communications theory course, similar to those subjects used in the experiment, were used for this pretest in December, 1985. They were administered the questionnaire after seeing all nine unmanipulated stories retained for the study.

7. To ensure the equivalence of the four groups, they were compared on their recall of information in the Louisiana Governor story. ANOVA results showed that there was no significant main effect for the captioning factor \( (F[1,96]=2.19, p=.142) \). There was a significant main effect, however, for the story type factor \( (F[1,96]=5.70, p=.019) \). This would have been a concern if hypotheses were being tested regarding the relative memorability of the two types of stories (picture vs. word). However, it was not considered a threat to validity in this case, given that the effects of captions were at issue, and their "relative" effect on the two types of stories (a possible interaction between the two main effects).

8. There were no appreciable differences between delayed and immediate visual recall, so only immediate visual recall scores were discussed further.
Table 2: Cell Means—Effects of Captioning on Recall from and Understanding of Picture Stories

<table>
<thead>
<tr>
<th>Story:</th>
<th>Hurricane</th>
<th>Earthquake</th>
<th>Terrorists</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Immediate Recall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.68</td>
<td>.81</td>
<td>.64</td>
<td>.71</td>
</tr>
<tr>
<td>No Caps</td>
<td>.71</td>
<td>.68</td>
<td>.67</td>
<td>.68</td>
</tr>
<tr>
<td><strong>Delayed Recall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.51</td>
<td>.81</td>
<td>.74</td>
<td>.69</td>
</tr>
<tr>
<td>No Caps</td>
<td>.51</td>
<td>.65</td>
<td>.74</td>
<td>.65</td>
</tr>
<tr>
<td><strong>Combined Recall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.60</td>
<td>.81*</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>No Caps</td>
<td>.61</td>
<td>.67</td>
<td>.71</td>
<td>.66</td>
</tr>
<tr>
<td><strong>Video Recall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.90</td>
<td>.85</td>
<td>.82*</td>
<td>.86</td>
</tr>
<tr>
<td>No Caps</td>
<td>.92</td>
<td>.89</td>
<td>.90</td>
<td>.90</td>
</tr>
<tr>
<td><strong>Understanding</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.71*</td>
<td>.33*</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td>No Caps</td>
<td>1.12</td>
<td>.65</td>
<td>.62</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the p=.05 level, two-tailed test

Note: Recall values are mean proportions of correct responses. Understanding values are means on three-point scale (0, 1, or 3).
Table 1: Cell Means—Effects of Captions on Recall from and Understanding of Word Stories.

<table>
<thead>
<tr>
<th>Story:</th>
<th>AIDS</th>
<th>Rainbow</th>
<th>Geneva</th>
<th>South</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Warrior</td>
<td>Talks</td>
<td>Africa</td>
<td></td>
</tr>
<tr>
<td>Immediate Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.73</td>
<td>.71**</td>
<td>.77*</td>
<td>.77</td>
<td>.75*</td>
</tr>
<tr>
<td>No Caps</td>
<td>.73</td>
<td>.46</td>
<td>.55</td>
<td>.71</td>
<td>.62</td>
</tr>
<tr>
<td>Delayed Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.73*</td>
<td>.58</td>
<td>.36</td>
<td>.35</td>
<td>.64*</td>
</tr>
<tr>
<td>No Caps</td>
<td>.57</td>
<td>.40</td>
<td>.37</td>
<td>.34</td>
<td>.55</td>
</tr>
<tr>
<td>Combined Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.73</td>
<td>.64*</td>
<td>.68</td>
<td>.72</td>
<td>.69*</td>
</tr>
<tr>
<td>No Caps</td>
<td>.65</td>
<td>.43</td>
<td>.61</td>
<td>.64</td>
<td>.59</td>
</tr>
<tr>
<td>Video Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.88</td>
<td>.89</td>
<td>.75</td>
<td>.92</td>
<td>.86</td>
</tr>
<tr>
<td>No Caps</td>
<td>.85</td>
<td>.85</td>
<td>.78</td>
<td>.90</td>
<td>.84</td>
</tr>
<tr>
<td>Understanding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caps</td>
<td>.73</td>
<td>.45</td>
<td>.45</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>No Caps</td>
<td>.57</td>
<td>.54</td>
<td>.43</td>
<td>.50</td>
<td></td>
</tr>
</tbody>
</table>

* Significant at the p<.05 level, two-tailed test
** Significant at the p<.01 level, two-tailed test

Note: Recall values are mean proportions of correct responses. Understanding values are means on three-point scale (0, 1, or 3).